

Ausschreibung Masterarbeit im Bereich Machine Learning / Text Mining

Der Lehrstuhl für Methoden der empirischen Bildungsforschung (Prof. Dr. Sven Hilbert) sucht für das laufende Sommersemester 2020 und/oder das kommende Wintersemester 2020/21 eine*n Studierende*n mit Interesse an einer Masterarbeit im Bereich Machine Learning / Text Mining.

Die Ausschreibung richtet sich an Studierende aller Fachrichtungen.

Die Forschung im Bereich der maschinellen Analyse und Verarbeitung von Text hat in den letzten Jahren immer mehr an Bedeutung gewonnen. Die dadurch gewonnen Erkenntnisse können zum Beispiel für Produktempfehlungen, Meinungs- und Verhaltensforschung, aber auch wissenschaftlich genutzt werden.

Inhaltlicher Fokus

In der Abschlussarbeit soll ein bestehender Datensatz (Twitter-Daten im Kontext von Bundesliga-Fußballspielen) mithilfe der Methoden des Text Mining ausgewertet werden.

Konkret ist das Ziel der Arbeit, die Performance verschiedener Arten des Feature Engineering in Hinblick auf die Verarbeitung von Kurznachrichtentexten (Tweets) auf Methoden des Machine Learning (z. B. Naive Bayes, Random Forest, etc.) zu untersuchen.

Voraussetzung für die Bearbeitung der Masterarbeit sind **Erfahrungen in mindestens einem der drei Bereiche** Machine Learning, Text Mining und der Programmiersprachen R und/oder Python.

Grundsätzlich sollte darüber hinaus die Bereitschaft bestehen,

- sich (weiter) in das Themengebiet des Text Mining einzuarbeiten, insbesondere in die verschiedenen Methoden, die in den einzelnen Phasen des Analyseprozesses sinnvoll angewendet werden können,
- sich (weiter) in die Statistiksoftware R einzuarbeiten, insbesondere in relevante Packages (z. B. StringR, tidyverse, kerasR),
- sich auch mit Techniken der Datenanalyse wie Decision Trees, Naive Bayes Classifier, Support Vector Machines, Deep Learning, Boosting und Resampling Methoden auseinanderzusetzen.

Ansprechpartnerin

Bei Interesse und/oder Fragen wenden Sie sich gerne an Elisabeth Kraus (elisabeth.kraus@ur.de).